Super-Resolution Image Reconstruction:

A Technical Overview

In most electronic imaging applications, images with high resolution (HR) are desired and often required. HR means that pixel density within an image is high, and therefore an HR image can offer more details that may be critical in various applications. For example, HR medical images are very helpful for a doctor to make a correct diagnosis. It may be easy to distinguish an object from similar ones using HR satellite images, and the performance of pattern recognition in computer vision can be improved if an HR image is provided. Since the 1970s, charge-coupled device (CCD) and CMOS image sensors have been widely used to capture digital images. Although these sensors are suitable for most imaging applications, the current resolution level and consumer price will not satisfy the future demand. For example, people want an inexpensive HR digital camera/camcorder or see the price gradually reduce, and scientists often need a very HR level close to that of an analog 35 mm film that has no visible artifacts when an image is magnified. Thus, finding a way to increase the current resolution level is needed.

The most direct solution to increase spatial resolution is to reduce the pixel size (i.e., increase the number of pixels per unit area) by sensor manufacturing techniques. As the pixel size decreases, however, the amount of light available also decreases. It generates shot noise that degrades the image quality severely. To reduce the pixel size without suffering the effects of shot noise, therefore, there exists the limitation of the pixel size reduction, and the optimally limited pixel size is estimated at about 40 μm^2 for a 0.35 μm CMOS process. The current image sensor technology has almost reached this level.

Another approach for enhancing the spatial resolution is to increase the chip size, which leads to an increase in capacitance [1]. Since large capacitance makes it difficult to speed up a charge transfer rate, this approach is not considered effective. The high cost for high precision optics and image sensors is also an important concern in many commercial applications regarding HR imaging. Therefore, a new approach toward increasing spatial resolution is required to overcome these limitations of the sensors and optics manufacturing technology.

One promising approach is to use signal processing techniques to obtain an HR image (or sequence) from observed multiple low-resolution (LR) images. Recently, such a resolution enhancement approach has been one of the most active research areas, and it is called super resolution (SR) (or HR) image reconstruction or simply resolution enhancement in the literature [1]-[61]. In this article, we use the term "SR image reconstruction" to refer to a signal processing approach toward resolution enhancement because the term "super" in "super

Sung Cheol Park, Min Kyu Park, and Moon Gi Kang

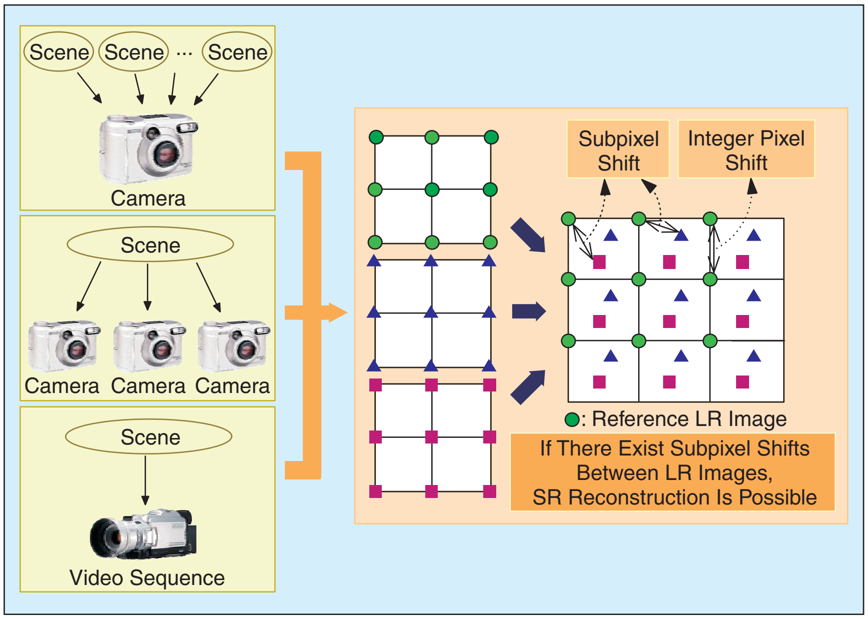
resolution" represents very well the characteristics of the technique overcoming the inherent resolution limitation of LR imaging systems. The major advantage of the signal processing approach is that it may cost less and the existing LR imaging systems can be still utilized. The SR image reconstruction is proved to be useful in many practical cases where multiple frames of the same scene can be obtained, including medical imaging, satellite imaging, and video applications. One application is to reconstruct a higher quality digital image from LR images obtained with an inexpensive LR camera/camcorder for printing or frame freeze purposes. Typically, with a camcorder, it is also possible to display enlarged frames successively. Synthetic zooming of region of interest (ROI) is another important application in surveillance, forensic, scientific, medical, and satellite imaging. For surveillance or forensic purposes, a digital video recorder (DVR) is currently replacing the CCTV system, and it is often needed to magnify objects in the scene such as the face of a criminal or the licence plate of a car. The SR technique is also useful in medical imaging such as computed tomography (CT) and magnetic resonance imaging (MRI) since the acquisition of multiple images is possible while the resolution quality is limited. In satellite imaging applications such as remote sensing and LANDSAT, several images of the same area are usually provided, and the SR technique to improve the resolution of target can be considered. Another application is conversion from an NTSC video signal to an HDTV signal since there is a clear and present need to display a SDTV signal on the HDTV without visual artifacts.

How can we obtain an HR image from multiple LR images? The basic premise for increasing the spatial resolution in SR techniques is the availability of multiple LR images captured from the same scene (see [4, chap. 4] for details). In SR, typically, the LR images represent different "looks" at the same scene. That is, LR images are subsampled (aliased) as well as shifted with subpixel precision. If the LR images are shifted by integer units, then each image contains the same information, and thus there is no new information that can be used to reconstruct an HR image. If the LR images have different subpixel shifts from each other and if aliasing is present, however, then each image cannot be obtained from the others. In this case, the new information contained in each LR image can be exploited to obtain an HR image. To obtain different looks at the same scene, some relative scene motions must exist from frame to frame via multiple scenes or video sequences. Multiple scenes can be obtained from one camera with several captures or from multiple cameras located in different positions. These scene motions can occur due to the controlled motions in imaging systems, e.g., images acquired from orbiting satellites. The same is true of uncontrolled motions, e.g., movement of local objects or vibrating imaging systems. If these scene motions are known or can be estimated within subpixel accuracy and if we combine these LR images, SR image reconstution is possible as illustrated in Figure 1.

In the process of recording a digital image, there is a natural loss of spatial resolution caused by the optical distortions (out of focus, diffraction limit, etc.), motion blur due to limited shutter speed, noise that occurs within the sensor or during transmission, and insufficient sensor density as shown in Figure 2. Thus, the recorded image usually suffers from blur, noise, and aliasing effects. Although the main concern of an SR algorithm is to reconstruct HR images from undersampled LR images, it covers image restoration techniques that produce high quality images from noisy, blurred images. Therefore, the goal of SR techniques is to restore an HR image from several degraded and aliased LR images.

A related problem to SR techniques is image restoration, which is a well-established area in image processing applications [62]-[63]. The goal of image restoration is to recover a degraded (e.g., blurred, noisy) image, but it does not change the size of image. In fact, restoration and SR reconstruction are closely related theoretically, and SR reconstruction can be considered as a second-generation problem of image restoration.

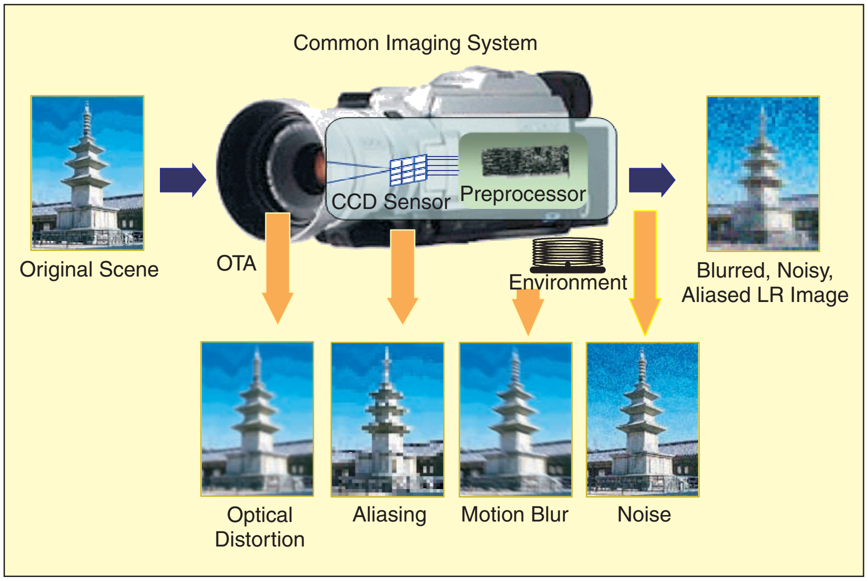
Another problem related to SR reconstruction is image interpolation that has been used to increase the size of a single image. Although this field has been extensively studied [64]-[66], the quality of an image magnified from an aliased LR image is inherently limited even though the ideal sinc basis function is employed. That is, single image interpolation



▲ *1. Basic premise for super resolution.*

cannot recover the high-frequency components lost or degraded during the LR sampling process. For this reason, image interpolation methods are not considered as SR techniques. To achieve further improvements in this field, the next step requires the utilization of multiple data sets in which additional data constraints from several observations of the same scene can be used. The fusion of information from various observations of the same scene allows us SR reconstruction of the scene.

The goal of this article is to introduce the concept of SR algorithms to readers who are unfamiliar with this area and to provide a review for experts. To this purpose, we present the technical review of various existing SR methodologies which are often employed. Before presenting the review of existing SR algorithms, we first model the LR image acquisition process.



▲ 2. *Common image acquisition system.*

Observation Model

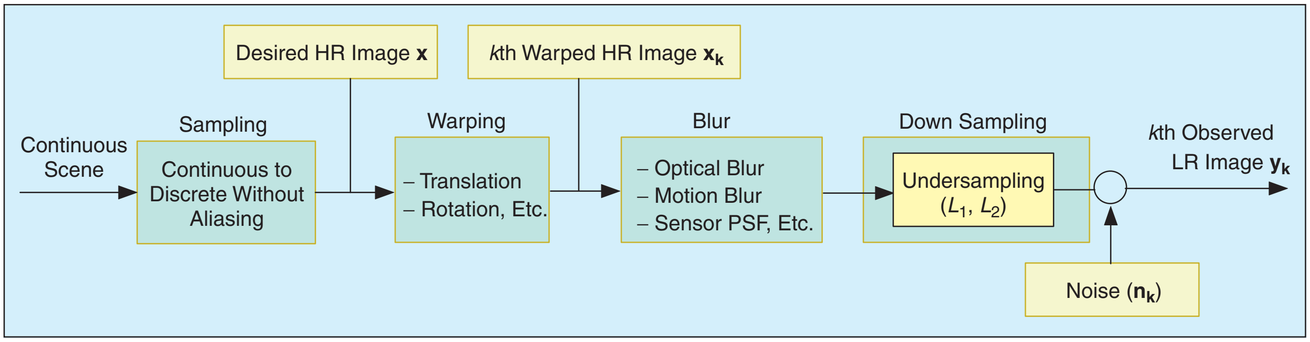
The first step to comprehensively analyze the SR image reconstruction problem is to formulate an observation model that relates the original HR image to the observed LR images. Several observation models have been proposed in the literature, and they can be broadly divided into the models for still images and for video sequence. To present a basic concept of SR reconstruction techniques, we employ the observation model for still images in this article, since it is rather straightforward to extend the still image model to the video sequence model.

Consider the desired HR image of size $L_1 N_1 \times L_2 N_2$ written in lexicographical notation as the vector $\mathbf{x} = [x_1, x_2, \ldots, x_N]^T$, where $N = L_1 N_1 \times L_2 N_2$. Namely, $\mathbf{x}$ is the ideal undegraded image that is sampled at or above the Nyquist rate from a continuous scene which is assumed to be bandlimited. Now, let the parameters L_1 and L_2 represent the down-sampling factors in the observation model for the horizontal and vertical directions, respectively. Thus, each observed LR image is of size $N_1 \times N_2$. Let the kth LR image be denoted in lexicographic notation as $\mathbf{y}_k = [y_{k,1}, y_{k,2}, \ldots, y_{k,M}]^T$, for $k = 1, 2, \ldots, p$ and $M = N_1 \times N_2$. Now, it is assumed that $\mathbf{x}$ remains constant during the acquisition of the multiple LR images, except for any motion and degradation allowed by the model. Therefore, the observed LR images result from warping, blurring, and subsampling operators performed on the HR image $\mathbf{x}$. Assuming that each LR image is corrupted by additive noise, we can then represent the observation model as [30], [48]

$$\mathbf{y}_k = \mathbf{D}\mathbf{B}_k\mathbf{M}_k\mathbf{x} + \mathbf{n}_k \quad \text{for} \quad 1 \le k \le p \tag{1}$$

where $\mathbf{M}_k$ is a warp matrix of size $L_1N_1L_2N_2 \times L_1N_1L_2N_2$, $\mathbf{B}_k$ represents a $L_1N_1L_2N_2 \times L_1N_1L_2N_2$ blur matrix, $\mathbf{D}$ is a $(N_1N_2)^2 \times L_1N_1L_2N_2$ subsampling matrix, and $\mathbf{n}_k$ represents a lexicographically ordered noise vector. A block diagram for the observation model is illustrated in Figure 3.

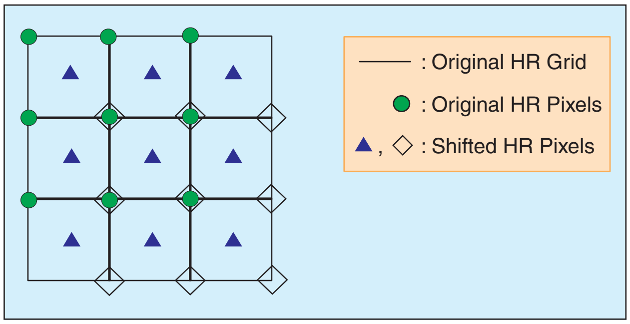
Let us consider the system matrix involved in (1). The motion that occurs during the image acquisition is represented by warp matrix $\mathbf{M}_k$. It may contain global or local translation, rotation, and so on. Since this information is



▲ 3. *Observation model relating LR images to HR images.*

The SR image reconstruction is proved to be useful in many practical cases where multiple frames of the same scene can be obtained, including medical imaging, satellite imaging, and video applications.

generally unknown, we need to estimate the scene motion for each frame with reference to one particular frame. The warping process performed on HR image **x** is actually defined in terms of LR pixel spacing when we estimate it. Thus, this step requires interpolation when the fractional unit of motion is not equal to the HR sensor grid. An example for global translation is shown in Figure 4. Here, a circle (○) represents the original (reference) HR image **x**, and a triangle (△) and a diamond (◇) are globally shifted versions of **x**. If the down-sampling factor is two, a diamond (◇) has (0.5, 0.5) subpixel shift for the horizontal and vertical directions and a triangle (△) has a shift which is less than (0.5,0.5). As shown in Figure 4, a diamond (◇) does not need interpolation, but a triangle (△) should be interpolated from **x** since it is not located on the HR grid. Although one could use ideal interpolation theoretically, in practice, simple methods such as zero-order hold or bilinear interpolation methods have been used in many literatures.



▲ *4. The necessity of interpolation in HR sensor grid.*

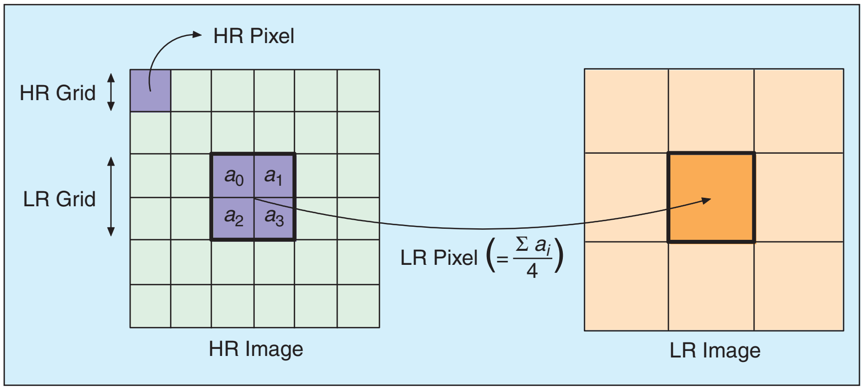
Blurring may be caused by an optical system (e.g., out of focus, diffraction limit, aberration, etc.), relative motion between the imaging system and the original scene, and the point spread function (PSF) of the LR sensor. It can be modeled as linear space invariant (LSI) or linear space variant (LSV), and its effects on HR images are represented by the matrix $\mathbf{B}_k$. In single image restoration applications, the optical or motion blur is usually considered. In the SR image reconstruction, however, the finiteness of a physical dimension in LR sensors is an important factor of blur. This LR sensor PSF is usually modeled as a spatial averaging operator as shown in Figure 5. In the use of SR reconstruction methods, the characteristics of the blur are assumed to be known. However, if it is difficult to obtain this information, blur identification should be incorporated into the reconstruction procedure.

The subsampling matrix **D** generates aliased LR images from the warped and blurred HR image. Although the size of LR images is the same here, in more general cases, we can address the different size of LR images by using a different subsampling matrix (e.g., $\mathbf{D}_k$). Although the blurring acts more or less as an anti-aliasing filter, in SR image reconstruction, it is assumed that aliasing is always present in LR images.

A slightly different LR image acquisition model can be derived by discretizing a continuous warped, blurred scene [24]-[28]. In this case, the observation model must include the fractional pixels at the border of the blur support. Although there are some different considerations between this model and the one in (1), these models can be unified in a simple matirx-vector form since the LR pixels are defined as a weighted sum of the related HR pixels with additive noise [18]. Therefore, we can express these models without loss of generality as follows:

$$\mathbf{y}_k = \mathbf{W}_k \mathbf{x} + \mathbf{n}_k, \quad \text{for} \quad k = 1, \ldots, p, \tag{2}$$

where matrix $\mathbf{W}_k$ of size $(N_1 N_2)^2 \times L_1 N_1 L_2 N_2$ represents, via blurring, motion, and subsampling, the contribution of HR pixels in **x** to the LR pixels in $\mathbf{y}_k$. Based on the observation model in (2), the aim of the SR image reconstruction is to estimate the HR image **x** from the LR images $\mathbf{y}_k$ for $k = 1, \ldots, p$.



▲ *5. Low-resolution sensor PSF.*

Most of the SR image reconstruction methods proposed in the literature consist of the three stages illustrated in Figure 6: registration, interpolation, and restoration (i.e., inverse procedure). These steps can be implemented separately or simultaneously according to the reconstruction methods adopted. The estimation of motion information is referred to as registration, and it is extensively studied in various fields of image processing [67]-[70]. In the registration stage, the relative shifts between LR images compared to the reference LR image are estimated with fractional pixel accuracy. Obviously, accurate subpixel motion estimation is a very important factor in the success of the SR image reconstruction algorithm. Since the shifts between LR images are arbitrary, the registered HR image will not always match up to a uniformly spaced HR grid. Thus, nonuniform interpolation is necessary to obtain a uniformly spaced HR image from a nonuniformly spaced composite of LR images. Finally, image restoration is applied to the upsampled image to remove blurring and noise.

The differences among the several proposed works are subject to what type of reconstruction method is employed, which observation model is assumed, in which particular domain (spatial or frequency) the algorithm is applied, what kind of methods is used to capture LR images, and so on. The technical report by Borman and Stevenson [2] provides a comprehensive and complete overview on the SR image reconstruction algorithms until around 1998, and a brief overview of the SR techniques appears in [3] and [4].

Based on the observation model in (2), existing SR algorithms are reviewed in the following sections. We first present a nonuniform interpolation approach that conveys an intuitive comprehension of the SR image reconstruction. Then, we explain a frequency domain approach that is helpful to see how to exploit the aliasing relationship between LR images. Next, we present deterministic and stochastic regularization approaches, the projection onto convex sets (POCS) approach, as well as other approaches. Finally, we discuss advanced issues to improve the performance of the SR algorithm.

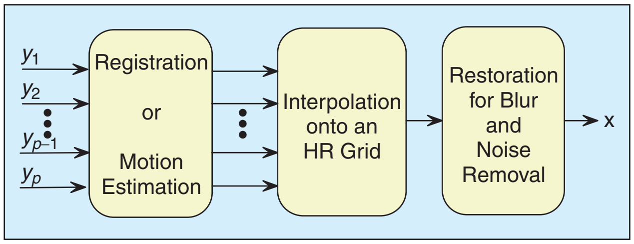
SR Image Reconstruction Algorithms

Nonuniform Interpolation Approach

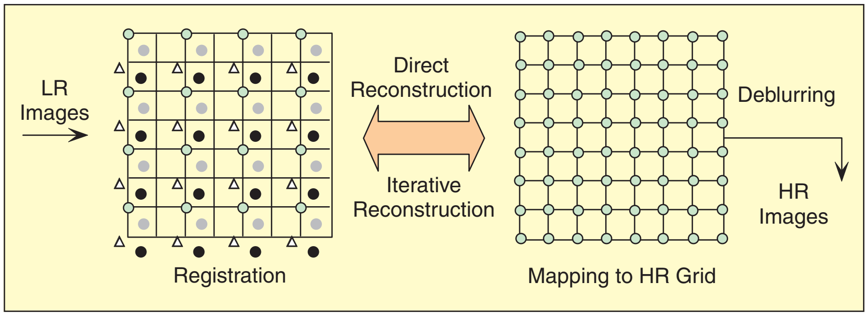
This approach is the most intuitive method for SR image reconstruction. The three stages presented in Figure 6 are performed successively in this approach: i) estimation of relative motion, i.e., registration (if the motion information is not known), ii) nonuniform interpolation to produce an improved resolution image, and iii) deblurring process (depending on the observation model). The pictorial example is shown in Figure 7. With the relative motion information estimated, the HR image on nonuniformly spaced sampling points is obtained. Then, the direct or iterative reconstruction procedure is followed to produce uniformly spaced sampling points [71]-[74]. Once an HR image is obtained by nonuniform interpolation, we address the restoration problem to remove blurring and noise. Restoration can be performed by applying any deconvolution method that considers the presence of noise.

The reconstruction results of this approach appear in Figure 8. In this simulation, four LR images are generated by a decimation factor of two in both the horizontal and vertical directions from the 256×256 HR image. Only sensor blur is considered here, and a 20-dB Gaussian noise is added to these LR images. In Figure 8, part (a) shows the image interpolated by the nearest neighborhood method from one LR observation, and part (b) is the image produced by bilinear interpolation; a nonuniformly interpolated image from four LR images appears in part (c), and a deblurred image using the Wiener restoration filter from part (c) is shown in part (d). As shown in Figure 8, significant improvement is observed in parts (c) and (d) when viewed in comparison with parts (a) and (b).

Ur and Gross [5] performed a nonuniform interpolation of an ensemble of spatially shifted LR images by utilizing the generalized multichannel sampling theorem of Papoulis [73] and Brown [74]. The interpolation is followed by a deblurring process, and the relative shifts are assumed to be known precisely here. Komatsu et al. [1] presented a scheme to acquire an improved resolution image by applying the Landweber algorithm [75] from multiple images taken simultaneously with multiple cameras. They employ the block-matching technique to measure relative shifts. If the cameras have the same aperture, however, it imposes severe limitations both in their arrangement and in the configuration of the scene. This difficulty was overcome by using multiple cameras with different apertures [6]. Hardie et al. developed a technique for real-time infrared image registration and SR reconstruction [7]. They utilized a gradient-based



▲ *6. Scheme for super resolution.*



▲ *7. Registration-interpolation-based reconstruction.*

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.